Research Article

Read counts from environmental DNA (eDNA) metabarcoding reflect fish abundance and biomass in drained ponds

Cristina Di Muri[1], **Lori Lawson Handley**[1], **Colin W. Bean**[2], **Jianlong Li**[3,4], **Graeme Peirson**[5], **Graham S. Sellers**[1], **Kerry Walsh**[6], **Hayley V. Watson**[1], **Ian J. Winfield**[7], **Bernd Hänfling**[1]

1 Evolution and Environmental Genomics Group (@EvoHull), School of Biological and Marine Sciences, University of Hull (UoH), Cottingham Road, Hull, HU6 7RX, UK

2 Scottish Natural Heritage, Caspian House, Mariner Court, Clydebank Business Park, Clydebank G81 2NR, UK

3 State Key Laboratory of Marine Resources Utilization in South China Sea, Hainan University, Hainan 570228, China

4 College of Marine Sciences, Hainan University, Haikou, Hainan 570228, China

5 Environment Agency, Mance House, Arthur Drive, Hoo Farm Industrial Estate, Worcester Road, Kidderminster DY11 7RA, UK

6 Environment Agency, Horizon House, Deanery Road, Bristol, BS1 5AH, UK

7 Lake Ecosystems Group, UK Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK

Corresponding author: Cristina Di Muri (c.di-muri@2016.hull.ac.uk)

Abstract

The sampling of environmental DNA (eDNA) coupled with cost-efficient and ever-advancing sequencing technology is propelling changes in biodiversity monitoring within aquatic ecosystems. Despite the increasing number of eDNA metabarcoding approaches, the ability to quantify species biomass and abundance in natural systems is still not fully understood. Previous studies have shown positive but sometimes weak correlations between abundance estimates from eDNA metabarcoding data and from conventional capture methods. As both methods have independent biases a lack of concordance is difficult to interpret. Here we tested whether read counts from eDNA metabarcoding provide accurate quantitative estimates of the absolute abundance of fish in holding ponds with known fish biomass and number of individuals. Environmental DNA samples were collected from two fishery ponds with high fish density and broad species diversity. In one pond, two different DNA capture strategies (on-site filtration with enclosed filters and three different preservation buffers versus lab filtration using open filters) were used to evaluate their performance in relation to fish community composition and biomass/abundance estimates. Fish species read counts were significantly correlated with both biomass and abundance, and this result, together with information on fish diversity, was repeatable when open or enclosed filters with different preservation buffers were used. This research demonstrates that eDNA metabarcoding provides accurate qualitative and quantitative information on fish communities in small ponds, and results are consistent between different methods of DNA capture. This method flexibility will be beneficial for future eDNA-based fish monitoring and their integration into fisheries management.

Key Words

Introduction

Environmental DNA (eDNA) metabarcoding is revolutionising biomonitoring in aquatic environments (Lawson Handley 2015; Harper et al. 2019; Jerde 2019; Sigsgaard et al. 2019). This approach relies on the molecular identification of organisms whose genetic material has been collected, isolated and extracted from water. Species identification occurs after PCR with broad-range primers followed by High Throughput Sequencing and matching

PENSOFT.

sequence reads against a reference database (see e.g. Valentini et al. 2016; Deiner et al. 2017 for an overview).

Environmental DNA metabarcoding has been recently suggested as a complementary biomonitoring strategy for the European Union Water Framework Directive (WFD, 2000/60/EC) which requires member states to assess the ecological status of freshwater bodies. Currently established WFD methodologies include the morphological identification and counting of phytoplankton, phytobenthos and benthic invertebrates or gillnetting and electrofishing for fish (Hering et al. 2018). Yet traditional biomonitoring methods have limitations which may hamper species' detectability or correct identification. They often lack broad applicability and they frequently impact on species' welfare, such as the use of gillnets for fish (Radinger et al. 2019). Environmental DNA metabarcoding has the advantage of detecting elusive and rare species, resolving cryptic species and identifying novel taxa through a non-invasive sampling approach (Blackman et al. 2017; Grey et al. 2018; Bylemans et al. 2019). The ease of eDNA collection also makes this approach suitable for remote location sampling, and the molecular identification of the genetic material does not require taxonomic expertise. Environmental DNA metabarcoding has been shown to outperform established methods for the assessment of freshwater fish community composition (Civade et al. 2016; Hänfling et al. 2016; Valentini et al. 2016; Pont et al. 2018; Sard et al. 2019).

The ability of eDNA metabarcoding to provide information on abundance and biomass is more controversial, and uncertainties regarding the quantitative power of eDNA metabarcoding are still present among the scientific community and monitoring agencies (Fonseca 2018; Lamb et al. 2019). This is particularly important given that species abundance is a crucial component of biodiversity surveillance and ecological monitoring schemes, and in view of ongoing biodiversity changes worldwide (Ficetola et al. 2018). Positive correlations between eDNA metabarcoding data (i.e. site occupancy or read counts) and fish abundance or biomass (as deduced by established surveys e.g. gill-netting) have been demonstrated in natural environments (Thomsen et al. 2012; Hänfling et al. 2016; Lawson Handley et al. 2019; Sard et al. 2019). However, estimates from established surveys also have their own biases and may not necessarily reflect true species abundance. Accurate data on organism-based measures of abundance from natural aquatic habitats are difficult to obtain without exhaustive sampling – such as draining down water bodies – and hence authentic comparisons with eDNA data in natural systems are, to our knowledge, still very rare.

A second key question is how replicable eDNA metabarcoding is with different field and laboratory protocols. Standardisation of protocols may overcome this issue, but a "one-size fits all" protocol would be unrealistic (Ruppert et al. 2019). For instance, eDNA capture methods are often chosen based on factors such as proximity/accessibility of sampling locations and the availability of lab equipment. At present, enclosed filters are usually preferred for on-site processing, especially when remote locations are sampled, and storage buffers are used for DNA preservation within the encapsulated filter (Spens et al. 2017; Li et al. 2018; Takahashi et al. 2020). For field workers this approach would be logistically simple, less prone to contamination and much easier to integrate into monitoring programmes compared to laboratory-based methods of eDNA capture. Open filter membranes allow a larger volume to be filtered, but suffer from field and transportation logistics, and are potentially more vulnerable to the risk of contamination (Li et al. 2018; Majaneva et al. 2018).

To evaluate the efficiency and suitability of different eDNA capture, a number of published studies have compared different approaches (precipitation versus filtration; on-site versus in laboratory), and a variety of filtration equipment, filters material and filters pore size (e.g. Deiner et al. 2015; Eichmiller et al. 2016; Lacoursière-Roussel et al. 2016; Minamoto et al. 2016; Djurhuus et al. 2017; Majaneva et al. 2018). Recent studies have also investigated the ability of different filter types (enclosed and open filters) and preservation methods (buffers and freezing) to provide quantitative estimates of eDNA using organisms' biomass and abundance estimates from artificial stocked ponds (Li et al. 2018) or from in-field visual surveys (Takahashi et al. 2020). Evaluation of the quantitative performance of filter types and preservation methods based on absolute values of species biomass and abundance in natural environments would greatly contribute to the implementation of future eDNA-based surveys.

In the present study we tested whether eDNA metabarcoding can provide accurate information on the community composition and fish biomass and abundance in ponds that were drained as part of an invasive species eradication programme. During the drain down, all fish were counted, measured and weighed, providing absolute measures of species abundance and biomass, and so avoiding the biases of established techniques used in previous studies. Secondly, we tested whether estimation of fish abundance and biomass with eDNA metabarcoding is consistent between different methods of DNA capture, by comparing Sterivex (hereafter also STX) enclosed filters preserved with three different buffers (ethanol, Longmire's solution and RNAlater) and open filtration (using Mixed Cellulose Ester; MCE filters and a vacuum pump) followed by freezing preservation at -20 °C.

Methods

Study site and collection of fish abundance and biomass data

The study was carried out at a UK fishing venue which consisted originally of three hydrologically-isolated stocked ponds (Upper, Middle and Lower Lake; Fig. 1A). This site was included in an Environment Agency (EA) eradication programme for non-native
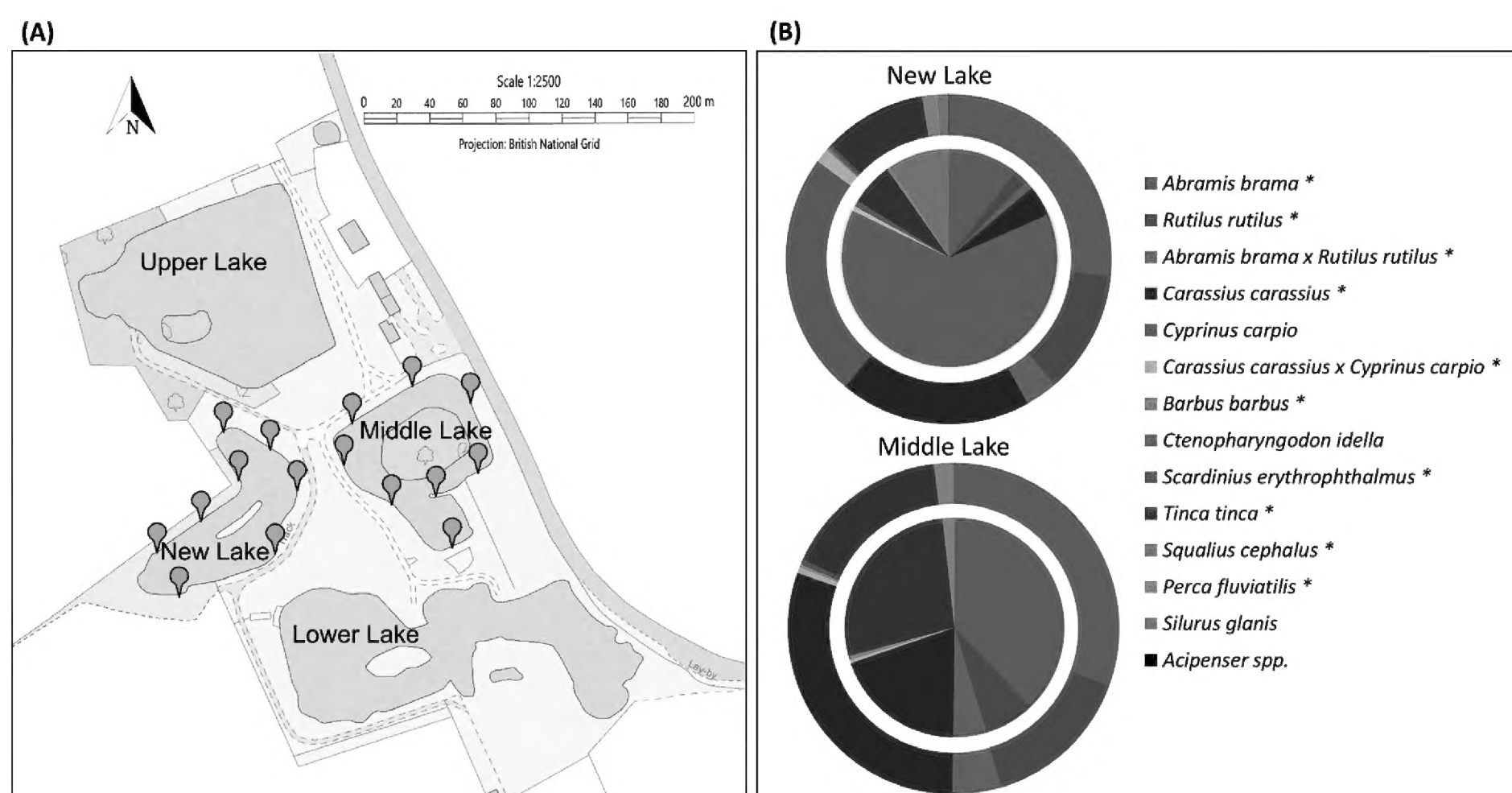
Figure 1. Map and fish diversity of the site surveyed. (**A**) Map of eDNA collection sites (in red) at the fishery venue. Map was downloaded and edited from Digimap (https://digimap.edina.ac.uk). (**B**) Fish species composition of the New Lake and Middle Lake after re-stocking (species with asterisk only). Ring pie charts (outer circles) show proportion of species composition by fish abundance (number of individuals); pie charts (inside circles) indicate proportion of species composition by fish biomass (kg).

topmouth gudgeon (*Pseudorasbora parva*), as part of a wider government strategy to tackle invasive species in the UK (GB Non-Native Species Secretariat, www.nonnativespecies.org). In November 2016, during the eradication programme, a new pond of 0.2 ha (hereafter "New Lake") was created, and the original three ponds drained. All fish over 150 mm total length from the original ponds were moved to the New Lake. During relocation, fish were individually checked for potentially hidden *P. parva* individuals in their gills and mouths. The original, empty ponds were partially refilled with water and treated with the piscicide rotenone by the EA to kill all potentially remaining specimens of *P. parva*. Original ponds were left fish-free for three months (from November 2016 to January 2017). On the 18th January 2017, New Lake was completely drained and all fish were moved back to the original ponds. During fish re-allocation, individual fish were morphologically identified by experts, counted and weighed, hence the exact fish biomass and population size could be calculated for each species and water body. Following the fishery owner's request, two of the original ponds (Upper and Lower Lake) became carp ponds, and they were re-stocked mainly with *Cyprinus carpio* and a few individuals of *Perca fluviatilis* and *Carassius carassius* x *C. carpio* hybrids. Middle Lake (0.3 ha) was re-stocked with 1,248 fish with a total biomass of 634.87 kg, equivalent to 2,116.23 kg/ha. The fish community included eight species and two hybrids with biomass and number of individuals ranging from 0.7 kg/1 individual (*Squalius cephalus*) to 240.6 kg/382 individuals (*Abramis brama*) (Fig. 1B; Suppl. material

1: Table S1). New Lake fish community was then calculated as the sum of fish species and hybrids counted and weighed after fish re-allocation to the original ponds, and included a total number of twelve species and two hybrids with biomass and numbers ranging from 0.7–1 kg/1 individual (*S. cephalus* and *Acipenser* spp.) to 1,715.2 kg/483 individuals (*C. carpio*) (Fig. 1B; Suppl. material 1: Table S1). Overall, New Lake contained 2,000 fish with a total biomass of 2,695.32 kg, equivalent to 13,476.6 kg/ha. Given the diverse fish communities of New Lake and Middle Lake, our eDNA metabarcoding analyses focused on these two ponds.

Water sample collection, filtration, and extraction

Water samples were taken on three separate occasions applying different strategies based on the goal of each occasion (see Fig. 2 for experimental design). New Lake was sampled the day before fish were transferred back to the original lakes (16th of January 2017) using MCE open filter membranes for eDNA capture (Fig. 2). We allowed one month after reintroductions for DNA dispersion in the water, and sampled Middle Lake on the 16th and 17th of February 2017, using replicated enclosed Sterivex filters and different preservation buffers (Middle Lake-STX; Fig. 2) and MCE open filter membranes (Middle Lake-MCE; Fig. 2).

All precautions to avoid contamination were taken while sampling. Fieldwork equipment was sterilised using 10% v/v chlorine-based commercial bleach (Elliott Hygiene Ltd, UK) and sterile gloves (STARLAB, UK) were changed at each sampling location. Blanks, consist-

ing of 2 L sampling bottles filled with ultra-purified water (Milli-Q), were included for each sampling occasion. Blanks were opened once in the field and then kept and processed alongside other water samples.

On each sampling occasion, eight 2 L water samples were collected equidistantly (~30 m apart) around the perimeter of each pond (Fig. 1A). Samples were collected by hand at the water surface by pooling five 400 mL sub-samples collected within a range of 5 m from the central location into a 2 L sterile plastic bottles (Gosselin Square HDPE, Fisher Scientific UK Ltd, UK). At each sampling occasion, immediately before filtration, a mixed sample was created using 200 mL aliquots from each of the eight water bottles collected in the field in order to evaluate differences of species detections with sampling strategies.

Samples for open filtration were placed inside cool boxes with ice packs, transported back to the laboratory and processed within 12 hours of collection. Environmental DNA was captured on 0.45 µm MCE membranes (47 mm diameter, Whatman, GE Healthcare) using a vacuum-pump and Nalgene™ filtration units. Filtration equipment was sterilised in 10% v/v chlorine-based commercial bleach (Elliott Hygiene Ltd, UK) for 10 min, then rinsed with 5% v/v MicroSol detergent (Anachem, UK) and with purified water. Filtration was stopped after 45 min and approximately 500 mL of water was filtered through each of two MCE open filter membranes per sample (i.e. 1 L of the 2 L total sample was filtered). Filter membranes were then stored in sterile 50 mm Petri dishes (Fisher Scientific UK Ltd, UK) sealed with parafilm (Bemis, Fisher Scientific UK Ltd, UK) and kept at -20 °C until DNA extraction.

Sterivex filtration was carried out in the field. Environmental DNA was captured using 0.45 µm Sterivex filter units (PVDF membrane, Merck Millipore) connected to a peristaltic pump (Easy Load II Peristaltic Pump, In-situ Europe Ltd, UK). On-site filtration was also carried out until an individual filter became clogged, otherwise it was stopped after 45 min. Approximately 350 mL were filtered through each Sterivex filter and three Sterivex units were used per sample. Each filter was then preserved using 2 mL of one of three different buffers: ethanol (≥ 99.5% v/v), Longmire's solution, and RNAlater.

All DNA extractions were carried out using the Mu-DNA protocol for water samples following adaptation for Sterivex as recommended in Sellers et al. (2018), and the DNA was eluted into 100 µL of TE buffer (see Supporting Information for details). Filter replicates of MCE open membranes from New Lake were co-extracted by placing both filters in a single tube for bead milling, whereas, to compare metabarcoding results of open membranes from the Middle Lake-MCE sampling, filter replicates were extracted separately (Fig. 2). For Sterivex units, DNA from buffers and filters was extracted separately as previous studies have shown that DNA can become suspended in the buffer (Spens et al. 2017; Fig. 2). After extractions, nucleic acid yield and purity were checked on a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific).

Contamination during laboratory procedures was minimised by using separated laboratories, located on different floors, for pre-PCR and post-PCR work. Pre-PCR procedures (DNA extraction and PCR preparation) were performed in a dedicated laboratory where only eDNA samples are handled. This laboratory has separated work stations for DNA extraction and PCR preparation. All equipment, instruments and benches are sterilised with 10% commercial bleach solution and 70% ethanol solution prior and after any work. PCR preparation occurred under UV-sterilised hoods with dedicated PCR pipettes.

Library preparation

Library preparation included a two-step PCR with a nested-tagging approach as described in Li et al. (2019a, b). In the first round of PCR, indexed primers targeting a 106 bp region within the mitochondrial 12S gene were used (Riaz et al. 2011; Kelly et al. 2014). The first round of PCRs was performed in a final reaction volume of 25 µL including 12.5 µL of Q5 Hot-Start High-Fidelity 2X Master Mix (New England Biolabs Inc., MA, USA), 1.5 µL of each indexed primer (10 µM; Integrated DNA Technologies, Belgium), 7.5 µL of molecular grade water (MGW; Fisher Scientific UK Ltd, UK) and 2 µL of DNA template at the original sample concentration. In order to avoid cross-contamination between samples as a consequence of evaporation and/or aerosols, reactions were prepared in 8-strip tubes with individually attached caps and covered with a drop of mineral oil (Sigma-Aldrich Company Ltd, UK). Amplifications were performed on Applied Biosystems Veriti thermal cyclers (Life Technologies, CA, USA) with the following conditions: initial denaturation at 98 °C for 5 min; 35 cycles of 98 °C for 10 sec, 58 °C for 20 sec and 72 °C for 30 sec; final elongation step at 72 °C for 7 min. Eighty-one samples, eight collection blanks, six PCR negatives (Molecular Grade Water, MGW), and four positives (genomic DNA [0.05 ng/ µL] from cichlid species not occurring in the UK, *Astotilapia calliptera* and *Maylandia zebra*) were amplified in triplicate. Amplicons were checked on 2% agarose gels stained with 10,000 × GelRed Nucleic Acid Gel Stain (Cambridge Bioscience, UK). Gels were imaged using Image Lab Software (Bio-Rad Laboratories Ltd, UK) to visually check for contamination in blanks/negatives, presence of target band and consistency of results among PCR replicates.

After visualisation, PCR triplicates were combined and samples belonging to the same collection site were pooled and normalised using different volumes as deduced from strength of PCR products on gels (no/ very faint band = 10 µL, faint band = 7.5 µL, bright band = 5 µL) using 1 µL of the positive samples and 5 µL of blanks/negatives for each pool (Alberdi et al. 2018).

Amplicon pools were cleaned using a double-size selection magnetic beads protocol (Bronner et al. 2013) with a ratio of 0.9X and 0.15X of magnetic beads (Mag-Bind RXNPure Plus, Omega Bio-tek Inc, GA, USA) to PCR products (detailed protocol in Supporting Informa-
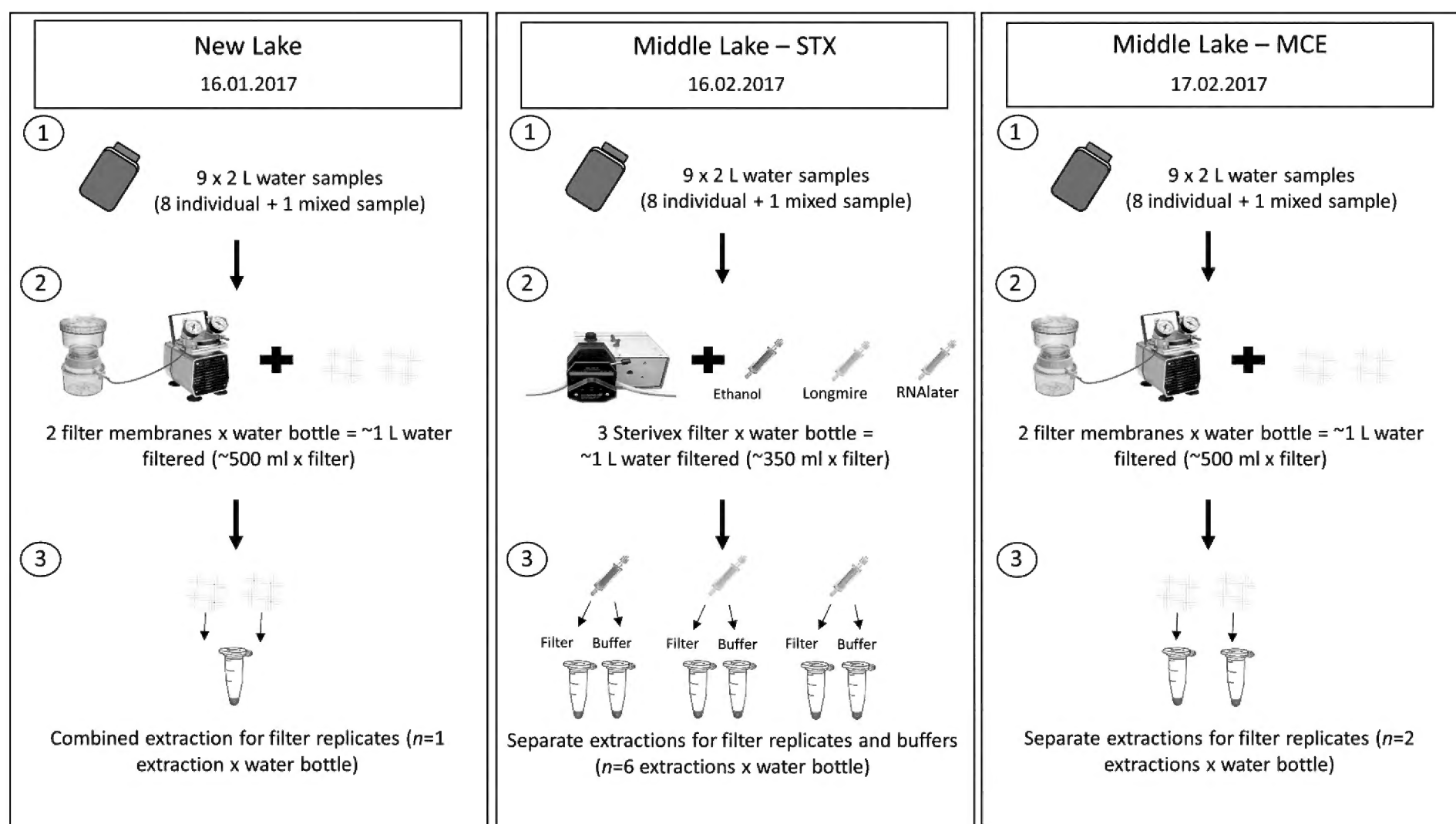
Figure 2. Experimental design. Panels show eDNA collection at different ponds (New Lake and Middle Lake) and processing strategies (Sterivex filters [STX] vs. Mixed Cellulose Ester open filters [MCE]). Numbers within the panels indicate the workflow from water sampling (1) to filtration (2) and DNA extraction (3).

tion). The double-size selection protocol and associated bead ratios were chosen to size-select the target amplicons from larger or smaller non-specific PCR products (e.g. primer dimers or non-target amplicons). Bead purification was followed by a second amplification where Illumina tags were added to each pool. Second PCRs were run in duplicate in a final reaction volume of 50 µL using 25 µL of Q5 Hot-Start High-Fidelity 2X Master Mix (New England Biolabs Inc., MA, USA), 3 µL of each Illumina tag (10 µM; Integrated DNA Technologies, Belgium), 15 µL of MGW (Fisher Scientific UK Ltd, UK) and 4 µL of pooled templates. PCRs consisted of: 95 °C for 3 min; 8 cycles of 98 °C for 20 sec and 72 °C for 1 min; and 72 °C for 5 min. PCR products were checked on a 2% agarose gel alongside their non-tagged products to check for size differences after tag addition and replicates were pooled. A second double-size selection bead purification was carried out with a ratio of 0.7X and 0.15X of magnetic beads/PCR products. Tagged amplicon pools were quantified using the Qubit 3.0 fluorometer and a Qubit dsDNA HS Assay Kit (Invitrogen, UK) and pooled with equimolar concentrations into a unique library. The final library was checked for size and integrity using the Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape (Agilent Technologies, CA, USA) and quantified using qPCR with the NEBNext Library Quant Kit for Illumina (New England Biolabs Inc., MA, USA). The fragment size of the final library was ~ 318 bp and, following qPCR, 13 pM library was loaded on the Illumina MiSeq with 10% PhiX using a 2 × 300 bp V3 chemistry (Illumina Inc., CA, USA).

Bioinformatics and statistical analyses

Raw sequencing data were demultiplexed using a custom Python script and subsequently analysed with metaBEAT (metaBarcoding and Environmental Analysis Tool) v0.97.11 (https://github.com/HullUni-bioinformatics/metaBEAT), an in-house developed pipeline. Quality trimming, merging, chimera detection, clustering and taxonomic assignment against a custom-curated 12S reference database (Hänfling et al. 2016) containing sequences for all UK freshwater fish species were performed. Unassigned sequences were subjected to a separate BLAST search against the complete NCBI nucleotide (nt) database at 100% identity. The number of reads assigned to fish species during the taxonomic assignment against the custom database (i.e. read counts) was used for downstream analyses in R v.3.5.1. (R Core Team 2018).

Total read count per sample was calculated as the sum of assigned and unassigned reads. The proportion of reads assigned to each fish species over the total read counts was then calculated on a sample by sample basis. A low-frequency noise threshold of 0.001 (0.1%) was applied across the dataset to reduce the probability of false positives arising from cross-contamination or tag-jumping (De Barba et al. 2014; Hänfling et al. 2016). Based on the level of contamination found in sampling/filtration blanks and PCR negatives, a second arbitrary threshold was applied and all records occurring with less than 50 reads assigned were removed.

Morphological identification of fish species revealed that a substantial amount of F1 hybrids (Fig. 1;

C. carassius x *C. carpio* and *A. brama* x *Rutilus rutilus*) were present. As community eDNA approaches are unable to differentiate hybrids from parental species these were grouped together for the purpose of our correlation analyses; i.e. data on biomass/abundance and eDNA read counts/site occupancy for hybrids and their parental species were pooled.

As read counts and site occupancy data were not normally distributed, Spearman's rank correlation coefficient was used to calculate correlations between biomass/abundance data and species average read counts and site occupancy for filter types and treatments. Graphs were plotted using *ggplot2* (Wickham 2016) and lines of best fit were drawn using the function *geom_smooth*. Spearman's correlation coefficients and significant levels were displayed using functions in *ggpubr* (Kassambara 2018). Species site occupancy was calculated as the number of filter replicates with positive detections over the total number of filter replicates collected and processed using the same treatment ($n = 8$).

VEGAN package v2.5-4 (Oksanen et al. 2019) was then used to test differences of fish communities between filter types (Sterivex and MCE membranes) and treatments (preservation buffers and freezing). *Betadisper* was used to investigate compositional variance of each group, and homogeneity of group dispersions was tested using ANOVA. Distances from the centroids of each treatment and the variance within treatment were visualised with a Principal Coordinates Analysis (PCoA). To test groups for compositional differences, a permutational multivariate analysis of variance (PERMANOVA), with replicates nested into each filter type, was carried out using the *adonis* function. Tests were performed on a square-root transformed abundance-weighed dissimilarity matrix (Bray-Curtis) of species composition.

Kernel density plots of fish species richness distribution across eDNA samples for each pond (New Lake and Middle Lake) and eDNA filtration/preservation strategy (Sterivex with buffers and MCE open filters replicates) were used to evaluate the number of fish species detected in the mixed samples compared to the mean species richness of eight individual samples. Density plots were built using the function *geom_density* implemented in *ggplot2* (Wickham 2016).

Lastly, sample-based species accumulation curves (SACs) were built using the function *specaccum* for each filter type and replicate.

Results

Sequencing outputs and bioinformatics

The total number of forward and reverse sequences across 98 samples (81 eDNA samples and 17 controls) was 10,751,170. Of these, 6,398,530 paired-end sequences passed the trimming quality filter and 92% were subsequently merged. 3,389,668 sequences remained after chimera detection and clustering with an average read count per sample of 40,042 (excluding control samples). Excluding the cichlid species used as positive controls, 16 Operational Taxonomic Units (OTUs), and 1,314,623 sequences were identified as fish taxa, with 100% match to the custom-curated fish reference database with thirteen OTUs remaining after applying the thresholds. All fish OTUs were identified to species level with the exceptions of records matching the family Percidae. Percidae records were manually assigned to *P. fluviatilis* as this was the only species of the family identified in the study area during fish relocation.

P. parva reads found in two Middle Lake-STX samples (279 and 148 reads) were also excluded from further analyses as after eradication this species was not physically present at the site surveyed.

### Environmental DNA metabarcoding fish diversity

OTUs from eleven of the twelve fish species translocated to New Lake were detected in eDNA samples, but two records were removed after applying thresholds. Sequences from the following taxa were detected at all eight sites within New Lake: *A. brama*, *C. carassius*, *C. carpio*, *P. fluviatilis*, *R. rutilus*, *Silurus glanis* and *Tinca tinca* (Fig. 4, S1) with *C. carpio* showing the highest read counts (about 40,000) and other species reads ranging from 1,831 of *S. glanis* to 23,618 of *A. brama* (Fig. 3, Suppl. material 1: Fig. S1). In addition, *Barbus barbus* was detected at two sites (202 reads), and *Ctenopharyngodon idella* at one site (71 reads) (Figs 3, 4, Suppl. material 1: Fig. S1). The presence of *Scardinius erythrophthalmus* was found at two sites with a low number of reads (38 and 25 reads) and, therefore, removed after applying the filter threshold (see metaBEAT raw data, Suppl. material 2: Table S2). Taxonomic assignment based on our reference database failed to detect *Acipenser* spp., yet 79 reads (at one site) matched the family Acipenseridae during the unassigned BLAST search against GenBank, however this record was excluded from further analyses (see unassigned blast data, Suppl. material 3: Table S3).

All nine possible OTUs corresponding to the species reintroduced were detected beyond threshold limits in Middle Lake in both sampling occasions (16th and 17th of February). Eight OTUs (*A. brama*, *R. rutilus*, *C. carassius*, *C. carpio*, *T. tinca*, *B. barbus*, *P. fluviatilis*, *S. cephalus*) were detected in both Middle Lake-STX and Middle Lake-MCE, and with all filter replicates (Figs 3, 4, Suppl. material 1: Fig. S1). Five of these fish OTUs (*A. brama*, *R. rutilus*, *C. carassius*, *C. carpio*, *T. tinca*) showed high site occupancy (all sites occupied) and number of reads assigned (Figs 3, 4, Suppl. material 1: Fig. S1). Detection was less consistent for one of the two least abundant species, *S. erythrophthalmus*. In Middle Lake-STX, *S. erythrophthalmus* was only detected in one filter replicates preserved with RNAlater (266 reads), and in Middle Lake-MCE, in filter membrane replicate 1 (333 reads; Figs 3, 4, Suppl. material 1: Fig. S1).
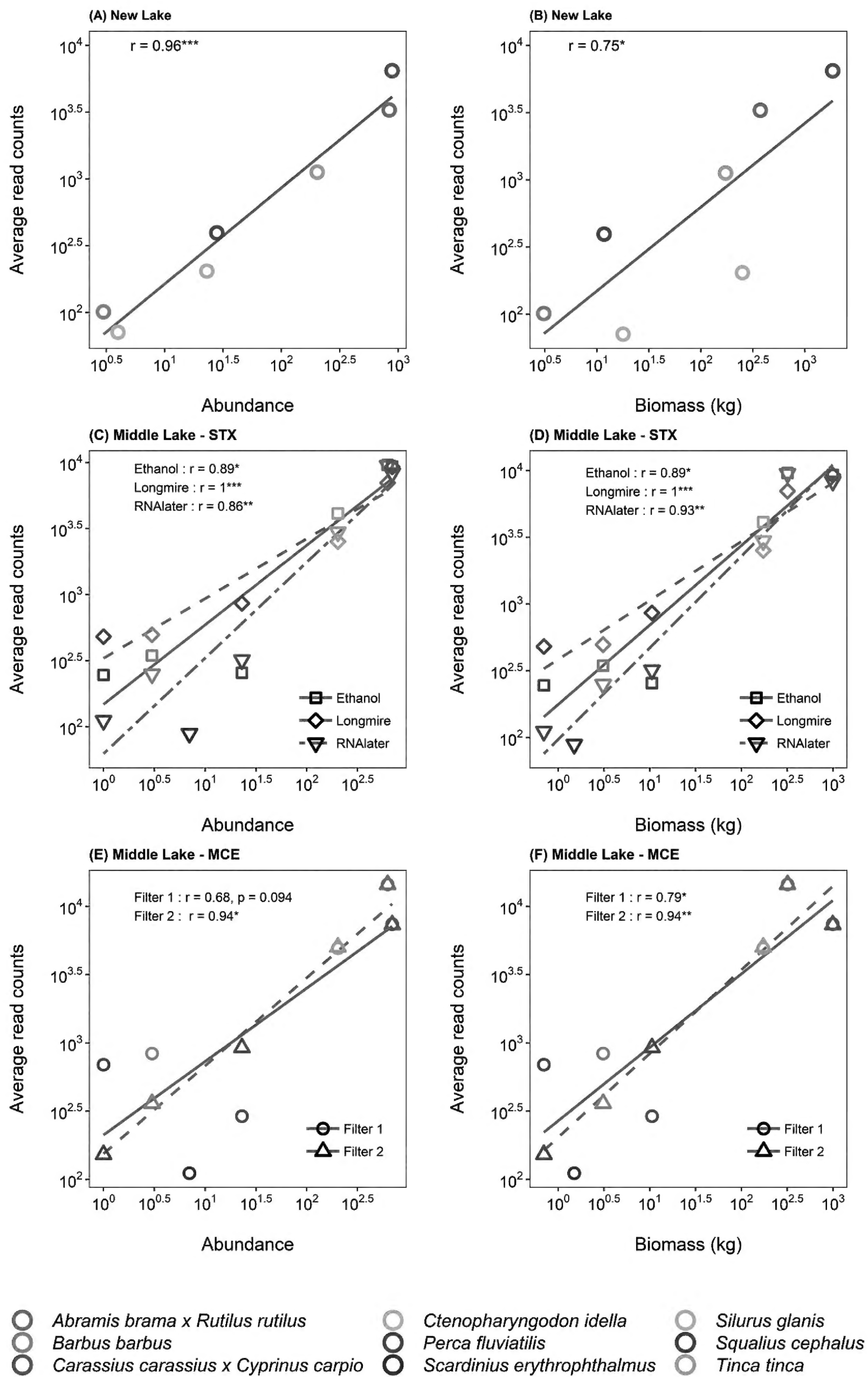
Figure 3. Correlations between eDNA metabarcoding read counts and fish abundance/biomass. Scatterplots showing lines of best fit and Spearman's correlations of fish species average read counts with abundance (number of individuals, on the left) and biomass (kg; on the right) at different sampling occasions. Panel (**A**) and (**B**) Spearman's correlations for New Lake; (**C**) and (**D**) Spearman's correlations for Middle Lake with Sterivex filters (STX); (**E**) and (**F**) Spearman's correlations for Middle Lake with open filter membranes (MCE). Plot axes were log transformed for better visualization. Significance codes: ***0.001; **0.01; *0.05.
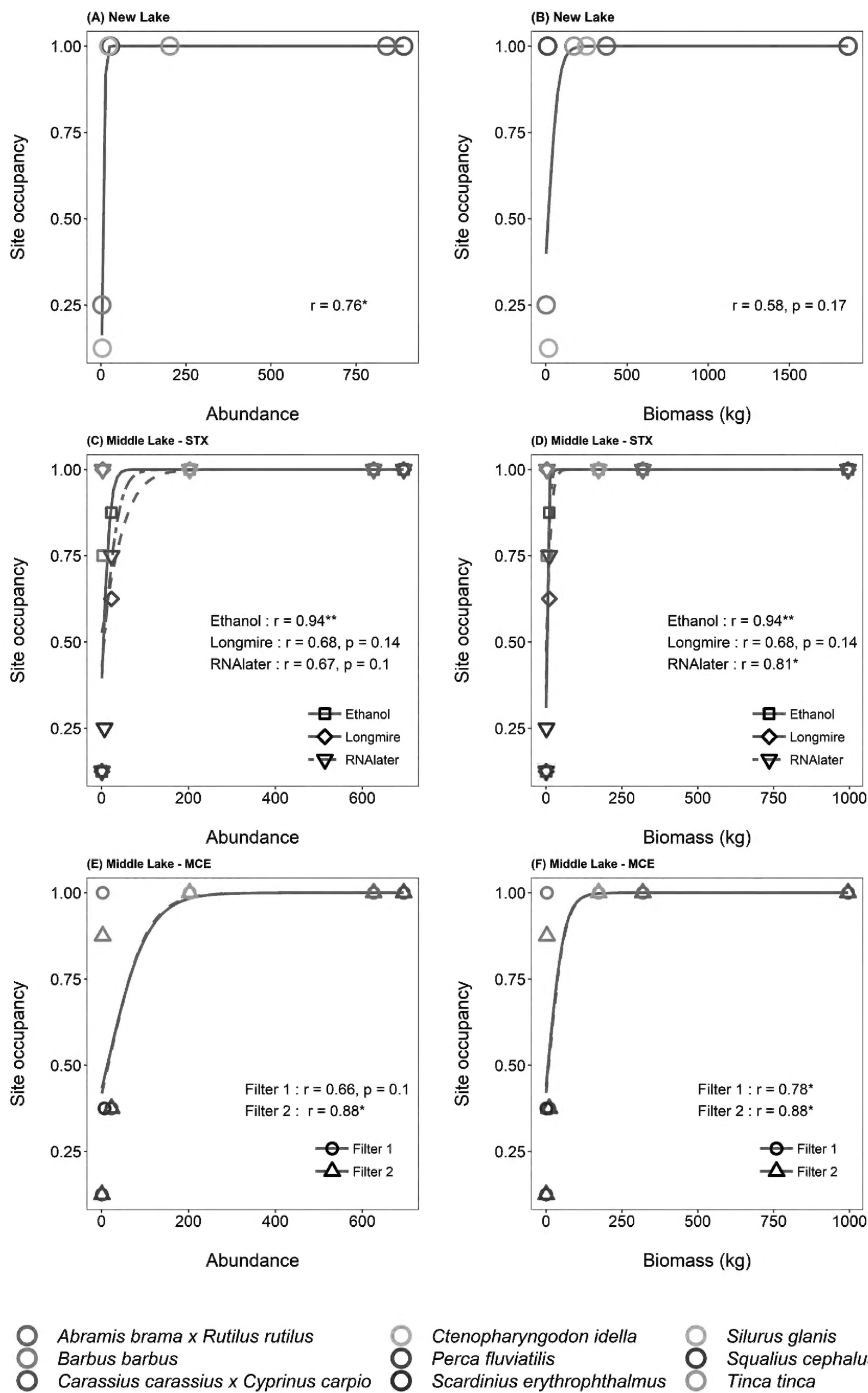
Figure 4. Correlations between eDNA metabarcoding site occupancy and fish abundance/biomass. Scatterplots showing lines of best fit and Spearman's correlations of fish species site occupancy with abundance (number of individuals, on the left) and biomass (kg; on the right) at different sampling occasions. Panel (**A**) and (**B**) Spearman's correlations for New Lake; (**C**) and (**D**) Spearman's correlations for Middle Lake with Sterivex filters (STX); (**E**) and (**F**) Spearman's correlations for Middle Lake with open filter membranes (MCE). Significance codes: ***0.001; **0.01; *0.05. Note: mixed samples were not included in the analyses.

Correlation between eDNA and biomass/abundance data

We evaluated the relationship between fish eDNA read counts/site occupancy of different filter replicates and fish biomass and abundance in New Lake and Middle Lake.

We observed a strong positive association between fish read counts and fish biomass (r = 0.75; p = 0.052; Fig. 3B) and especially between read counts and abundance (r = 0.96; p < 0.001; Fig. 3A) for samples collected from New Lake.

Spearman's correlations were calculated separately for each filter type (Sterivex/filter membranes) and filter replicate for samples collected from Middle Lake (Middle Lake-STX, Middle Lake-MCE). Fish read counts for all replicates and filters were positively correlated to both fish biomass and abundance. The highest associations were observed when read counts of Sterivex filter replicates were compared with biomass (Ethanol: r = 0.89, p = 0.019; Longmire: r = 1, p < 0.001; RNAlater: r = 0.93, p = 0.0025; Fig. 3D), and abundance (Ethanol: r = 0.89, p = 0.019; Longmire: r = 1, p < 0.001; RNAlater: r = 0.86, p = 0.014; Fig. 3C).

For MCE open filter membranes (Middle Lake-MCE), there was a significant correlation between read counts and biomass for both filter replicates (r = 0.79, p = 0.036; r = 0.94, p = 0.0048; Fig. 3F) and between read count and abundance for filter 2 (r = 0.94, p = 0.048; Fig. 3E), but the correlation between read count and abundance was not significant for filter 1 (r = 0.68, p = 0.094; Fig. 3E).

Positive but weaker correlations of New Lake eDNA samples were observed when species site occupancy was associated with fish biomass (r = 0.58, p = 0.17; Fig. 4B) and abundance (r = 0.76, p = 0.049; Fig. 4A).

Fish site occupancy of Middle Lake filter replicates (Middle Lake-STX, Middle Lake-MCE) was also positively correlated to both fish biomass and abundance with, however, weaker associations. Correlation coefficients and significance of the Spearman's correlations varied between filter replicates of both filter types. The strongest associations were observed when site occupancy of Sterivex filters preserved with ethanol were correlated with abundance and biomass (Ethanol: r = 0.94, p = 0.0051; Fig. 4C, D), but also when site occupancy of MCE open filter membrane replicate 2 were associated with fish species biomass and abundance (r = 0.88; p = 0.021; Fig. 4E, F).

Effect of sampling and filtration strategies on fish community eDNA data

To evaluate the effect of different sampling strategies the mean species richness of individual samples was compared to the species richness of the mixed sample at each sampling occasion and treatment (Fig. 5A). Overall, the number of fish species detected in the mixed samples was very close, and most of the time higher, than the average number of species detected in indi-

vidual field samples with the only exception of MCE filter replicate 2 (Fig. 5A). The fish species not represented in the mixed samples were often the low-occurrence taxa of the sites surveyed, and generally, excluding MCE filter 2, a number of two fish species were missing in the mixed samples. For example, in the New Lake mixed sample *B. barbus* and *C. idella* were not detected. *S. cephalus* and S. e*rythrophthalmus* were not represented in Middle Lake-STX (ethanol, RNAlater and Longmire's preservation) nor in Middle Lake-MCE filter 1 and 2 with the latter one additionally missing *B. barbus* and *P. fluviatilis*.

There were no differences between fish community composition of different filter types (ANOVA $F = 0.8521$, $p = 0.3611$; Fig. 5B) or filter replicates (ANOVA $F = 0.6495$, $p = 0.6305$; Fig. 5B).

There was no significant difference between centroids of Middle Lake fish communities described by eDNA metabarcoding when using different filter types (PERMANOVA; $R^2 = 0.23278$; $p = 0.7231$) or different preservation methods (buffers and freezing; $R^2 = 0.03795$; $p = 0.7231$). However, more variation (23%) was explained by the use of different DNA capture methods (MCE versus Sterivex), compared to within filter treatment (3.8%).

Species accumulation curves of both Sterivex and MCE filters showed that approximately six samples are required to detect all fish species when filter replicates are combined (Fig. 5C). SACs of single filter replicates for Sterivex filters showed higher rates of species detection with RNAlater preservation compared to Longmire's or ethanol preservation (Fig. 5C). For the MCE open filters, most of the species were recovered with the first filter membrane, with only a slight improvement in detection rate when the second membrane was included (Fig. 5C).

Discussion

With the advent of the next-generation eDNA-based monitoring surveys there is a growing interest in whether eDNA metabarcoding can generate accurate semi-quantitative data. Previous studies in natural environments have focussed on indirect estimates of fish abundance from established surveys which have their own inherent biases. Here, we used absolute data on fish abundance and biomass from drained ponds and found that read counts from eDNA metabarcoding consistently correlate with both fish abundance and biomass. Moreover, the present study suggests that the use of different eDNA capture (Sterivex vs. MCE open filters) and storage methods (buffers and freezing) produce repeatable results of fish diversity, composition and biomass/abundance estimates. We additionally show that the collection of spatial and filter replicates enhances species detection probability for rare species, thus sample coverage and replication are an important consideration in experimental design.
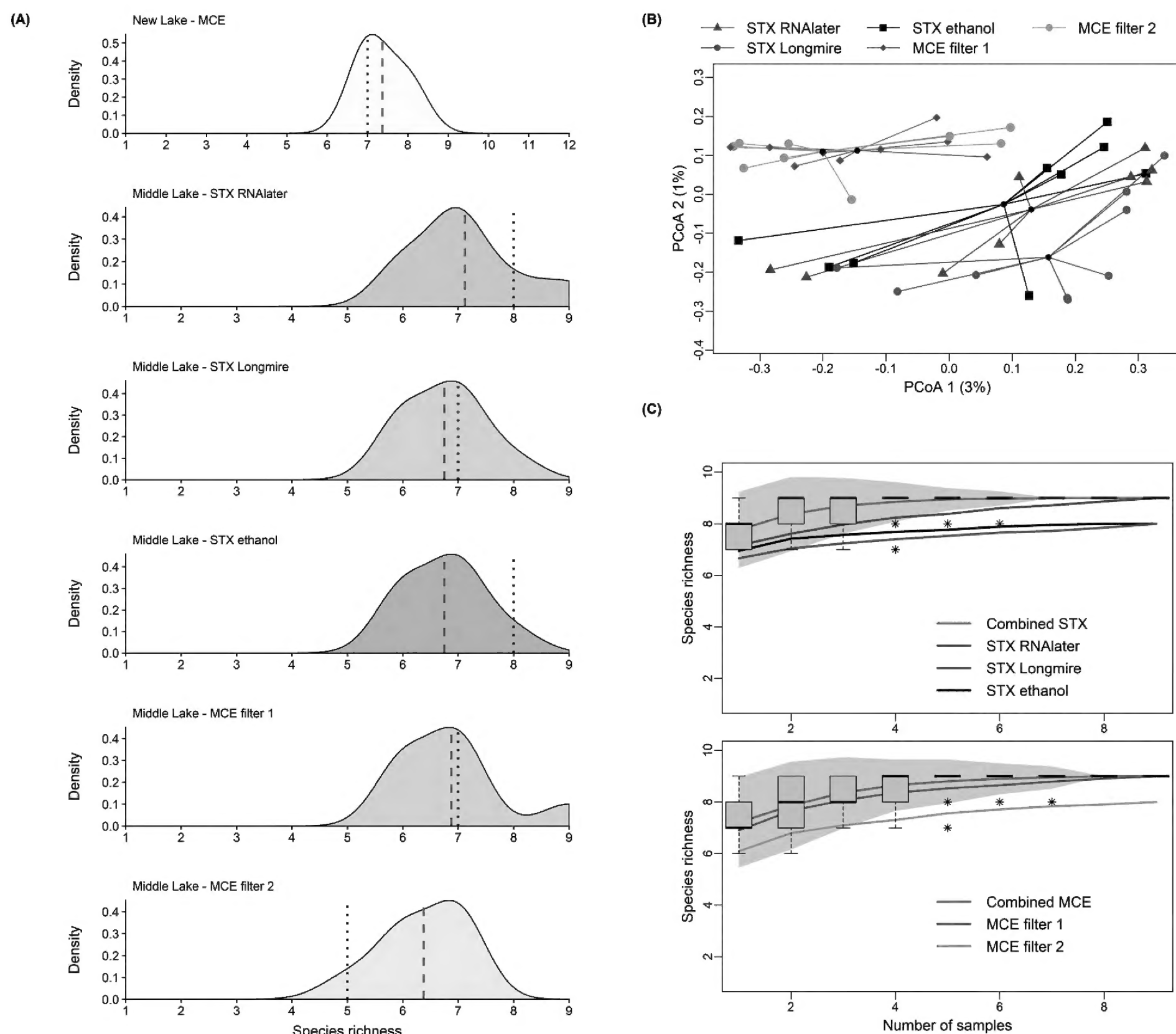
Figure 5. Environmental DNA metabarcoding fish community plots for different filter types and treatments. (**A**) Kernel density plots showing distribution of species richness across eDNA samples collected from different ponds and with different filtration and filter preservation strategies. The dashed blue lines indicate the mean species richness of individual eDNA samples ($n = 8$), the dotted black lines indicate the species richness of mixed samples (pooled aliquots of individual samples) at each sampling occasion and filtration/preservation strategy. The x axes represent the fish species richness in each pond surveyed (New Lake = 12 species; Middle Lake = 9 species) (**B**) PCoA plot showing distances from centroids of filter types (MCE and Sterivex [STX]; ANOVA $F = 0.8521$, $p = 0.3611$) and treatments (buffers and freezing; ANOVA $F = 0.6495$, $p = 0.6305$). Distances from centroids were calculated upon a dissimilarity matrix (Bray-Curtis) of fish species read counts. (**C**) Species accumulation curves for filter replicates of Sterivex filters preserved with buffers (top) and MCE filters with freezing preservation (bottom). In both figures, golden curves are calculated based on the sum of species when filter replicates/treatments of the same filter type are combined. 95% confidence intervals refer to the golden curves and boxplots of these curves show distribution of species diversity as inferred from the method "random", which add sites in random order and was used for the SACs. Asterisks represent outliers.

Fish species detection

All 12 fish taxa were successfully detected in both fishery ponds surveyed with the only exception of *S. cephalus* in New Lake (single specimen of 0.7 kg; Fig. 1B, Suppl. material 1: Table S1). Our findings are in line with other studies corroborating the ability of eDNA metabarcoding to describe fish diversity in lentic environments (Civade et al. 2016; Hänfling et al. 2016; Li et al. 2018; Lawson Handley et al. 2019; Zhang et al. 2020). The appropriate sam-

pling effort, such as volume of water and spatial replicates collected, may vary according to the waterbody features (i.e. surface area, depth, heterogeneity) and other environmental and biological factors (Civade et al. 2016; de Souza et al. 2016; Lawson Handley et al. 2019). In this study, the collection of eight, evenly distributed 2 L water samples from the ponds' shore and the filtration of ~ 1 L from each water bottle provided sufficient coverage of the fish community of the ponds surveyed. An average of 1 L water filtered from each of the eight 2 L samples collected from the

edge of Middle Lake was appropriate for species detection at both sampling occasions and for the different filter types used. In fact, rarefaction curves (Fig. 5C) demonstrated that when filter replicates are combined, six 2 L water samples (with ~1 L of water filtered) are sufficient to unveil the total fish composition of this intensively stocked (0.3 ha and 2,116.23 kg/ha of fish density; Fig. 5A, C). In line with other eDNA metabarcoding studies, we suggest that near-shore sampling provides adequate species coverage as previously observed in larger and deeper lentic environments with complex fish species assemblages where a greater number of species has been detected inshore as opposed to offshore waters (Hänfling et al. 2016; Zhang et al. 2020). Here, we additionally highlight that an adequate sampling effort is paramount for describing species occurrence within a water body. In small, shallow lentic systems eDNA is thought to be homogeneously distributed in the water (Thomsen et al. 2012; Evans et al. 2017) even though the signal strength may increase closer to its source (Li et al. 2019b). Yet, we demonstrated that eDNA concentration of low-abundant species DNA is very localised, hence intensive sampling efforts and the collection of an adequate number of replicates is required to detect low-occurrence taxa. For example, our mixed samples (pooled water aliquots of field individual samples) consistently detected the common fish species at all sampling occasions, but failed to detect individuals or low-abundance taxa. Spatial pooling is therefore inefficient for detailed biodiversity surveillance as suggested by Zhang et al. (2020), who, on a larger spatial scale with higher number of PCR replicates, still found reduced OTUs detection in mixed water samples. In line with these results, we demonstrated that eDNA detection rate is enhanced with spatial and technical replication as well as with the increased water volume filtered.

Of particular interest is the detection of *P. parva* DNA in Middle Lake samples as this invasive species was the target of the eradication programme and present in extremely high abundance before the ponds were drained and treated with a piscicide. The persistence of *P. parva* as living organisms within the pond appears extremely unlikely due to the effective eradication methods used in combination with the relatively small size of the water body (Britton et al. 2008; Genovesi and Carnevali 2011). Furthermore, *P. parva* has not been recorded in these lakes since the eradication programme. Contamination could have occurred during the water sampling or in the laboratory resulting in false positive detection (e.g. Hänfling et al. 2016). However, no *P. parva* DNA was found in any of the control samples or in the water samples from New Lake. A possible explanation for this record is that *P. parva* eDNA originated from sediment re-suspension in the water column during sampling or from carcasses remaining at the bottom of the pond. Environmental DNA is known to be less concentrated and less persistent in water compared to sediment, where it remains detectable for over three months also when species are removed from the system (Turner et al. 2015). A further reasonable option would be to consider this result as a true record even if we have no evidence that the species re-colonised the pond after the eradication. Previous studies have suggested that *P. parva* may suffer from recruitment failure and local extirpation when population numbers are low due to human or natural disturbance (Copp et al. 2007; Davidson et al. 2017). Therefore, when monitoring the success of eradication attempts, extreme caution should be taken with false-positive or false-negative detections for the target species and the use of conventional methods to corroborate eDNA detections has been recommended (Davison et al. 2019; Robinson et al. 2019).

Read counts correlate to biomass and abundance

To our knowledge, this is the first published study to date where the correlation between eDNA metabarcoding data and actual measures of species biomass and abundance in semi-natural lentic systems has been investigated. Takahara et al. (2012) previously demonstrated that eDNA concentrations in ponds artificially stocked with common carps reflected the species absolute biomass. Using a similar environmental setting, we found that our eDNA metabarcoding results accurately reflect abundance patterns and reveal positive and mostly significant correlations between read counts and fish species biomass (weight) and abundance (Fig. 3). Recently Kelly et al. (2019) demonstrated that when amplification efficiency is high in PCR-based studies, proportional indices of eDNA reads capture trends in taxon biomass with high accuracy. Our study supports these results as we found that the species' read counts were an accurate quantitative parameter to describe taxon biomass and abundance.

Positive associations were observed between species site occupancy and fish biomass/abundance, however, less significant than correlations with read counts (Fig. 4). In our study system, the relatively small size of the water bodies surveyed, coupled with the high fish densities, resulted in relatively homogeneous distribution of the common species' eDNA, generating a better representation of fish biomass and abundance when read counts were used for quantitative inferences. In larger and heterogeneous lentic environments, the spatial variation of the species' eDNA signal is likely to be as or more reliable than read counts for quantitative estimates (Hänfling et al. 2016; Lawson Handley et al. 2019; Sard et al. 2019).

Current uncertainties regarding the quantitative power of eDNA metabarcoding ultimately originate from our lack of knowledge on the origin and fate of eDNA in aquatic systems (Klymus et al. 2015; Lacoursière-Roussel et al. 2016; Sassoubre et al. 2016). Age, physiology, life history and metabolic rate all play a role in the amount of eDNA released (eDNA shedding rate) from organisms into their surroundings (Barnes et al. 2014; Goldberg et al. 2016; Ruppert et al. 2019). Physical, chemical and biological forces such as dilution, sedimentation and re-suspension, hydrolysis, oxidation and microbial activity, can all influence eDNA persistence and dynamics within aquatic habitats (Turner et al. 2015). In addition, the degradation

of genetic material is also promoted by high temperature and acidity (Seymour et al. 2018; Ruppert et al. 2019). In our study system, the fish age distribution was relatively narrow, therefore, reducing the effect of different eDNA shedding rates from distinct life stages and age classes. Moreover, the ponds surveyed were similar in terms of high fish density and species composition and were also exposed to stable environmental conditions that positively influenced the reproducible and reliable quantitative characterisation of the fish communities investigated.

A lack of robust sampling and metabarcoding protocols may also contribute to a distortion of the observed diversity patterns. Insufficient sampling effort, inhibition, primer biases, sequencing artefacts, database inaccuracy and contamination are the main methodological sources of bias (McKee et al. 2015; Grey et al. 2018; Collins et al. 2019; Wood et al. 2019). In the present study, the quantitative fish assessment of the two ponds surveyed demonstrates the accuracy of optimised eDNA metabarcoding protocols to reflect species biomass and abundance. In recent years, sampling, laboratory, and bioinformatics workflows have been progressively refined for the characterisation of fish communities within UK freshwater ecosystems (Hänfling et al. 2016; Li et al. 2018; Lawson Handley et al. 2019; Li et al. 2019a). Here, we have demonstrated that optimised sampling strategy, enhanced extraction protocol with an additional inhibitor removal step (Sellers et al. 2018), replication during PCRs and development of a custom-curated database with new reference sequences, strengthened the probability of detection, reduce taxonomic assignment bias, and overall provided reliable quantitative data of fish biomass and abundance.

Suitable eDNA metabarcoding data for quantitative fish monitoring require comparable measures of biomass and abundance across studies and over time, for example, to detect trends in abundance of fish populations. In light of this, the use of site occupancy appears a more practical approach as abundance/biomass estimates from site occupancy are easily comparable even across studies with uneven sampling efforts (Li et al. 2019a). Conversely, quantitative estimates from eDNA metabarcoding read counts will need to be adjusted to standardised metrics. One possible approach, proposed by Ushio et al. (2018), is the use of internal standards to convert raw sequence reads from metabarcoding into standardised reads using standard curves of known copy number.

Impact of DNA capture and preservation methods

In our study the correlations between sequence read counts and species abundance/biomass were consistently high for all filtration treatments with average correlation coefficients of 0.93 for Sterivex filters and 0.84 for MCE filters (Fig. 3). The variation of correlations observed between filter types may be explained by differences between read counts assigned to species as a result of different water volumes filtered between Sterivex and MCE filters (Suppl. material 1: Fig. S1). However, for equal amounts of water filtered and high DNA concen-

trations, MCE open filter membranes usually capture a higher amount of DNA compared to enclosed Sterivex filters possibly due to the tendency of Sterivex units to clog more easily (Li et al. 2018; Takahashi et al. 2020). Quantitative differences between filter types may also vary with the target species as observed in this study (Takahashi et al. 2020). In fact, while we observed a general trend of higher species read counts in MCE filters, we also observed the opposite trend for *C. carpio* which showed lower reads in MCE filter replicates compared to Sterivex filters (Suppl. material 1: Fig. S1).

The higher species richness found in Sterivex filters preserved with RNAlater and open filter membrane replicate 1 resulted from the detection of only one low-abundant taxon within the pond (*S. erythropthalmus*; Fig. 1). We therefore consider this result a stochastic effect between filter replicates or storage methods.

Overall, we found that both filter types showed a good representation of fish diversity and community composition and, consequently, we suggest that they can be used interchangeably depending on time, resources and location of the study. Sterivex filters, for instance, are effective for field processing of water samples, facilitating collection in remote locations. After sample collection, Sterivex are immediately filtered on-site (using peristaltic pumps or sterile syringes) and the risk of contamination is reduced because of the lack of filter handling (Spens et al. 2017; Li et al. 2018). In the present study, there was no evidence of higher contamination in MCE open filter membranes compared to Sterivex filters indicating that preventing on-site and in-lab contaminations is sufficient to minimise/avoid DNA contaminations regardless of the filter types' choice. The use of Sterivex filters, or enclosed filters in general, is however more amenable to large-scale monitoring programmes for environmental managers or citizen science projects (Biggs et al. 2015; Buxton et al. 2018; Larson et al. 2020). Nevertheless, Sterivex filters are currently almost 15 times more expensive than MCE open filters, DNA extraction is more time-consuming, and, when syringes are used for filtration, the Sterivex method requires a large amount of disposable plastic consumables. The use of prepacked sterile syringes is nonetheless preferred over pumps' suction (vacuum or peristaltic) to reduce filtration time (this paper; Li et al. 2018).

Conclusion

This study underpins valuable considerations for the quantitative estimates of eDNA metabarcoding data. We demonstrated that eDNA metabarcoding data correlate with actual abundance and biomass of fish communities within small freshwater systems with high fish density.

Established methods (i.e. hydroacoustic, electrofishing, gillnetting) for obtaining quantitative estimates of fish abundance are resource intensive and may not be suitable for all water bodies and species (Winfield et al. 2009). Furthermore, quantitative interpretation of data is often complex (hydroacoustic) or relies on large sampling effort (netting/electrofishing) (Winfield et al. 2009), hence

becoming costly in terms of financial, human resources and habitat disturbance or species mortality. More importantly, these methods can be prone to errors as they are not exhaustive sampling methods and, therefore, can only provide approximation of species abundance.

Environmental DNA metabarcoding is arguably a more flexible tool, adaptable to all aquatic environments and fish species, is non-lethal, and the sources of errors can be minimised through a careful optimisation of field and laboratory protocols.

Monitoring trends in population size and community structure is paramount to the assessment of species health and viability, and the outputs are required to undertake management actions and to guide conservation decisions (Kull et al. 2008). Implementation of eDNA metabarcoding will drive a step-change towards non-invasive monitoring strategies for next-generation ecosystems surveillance. Environmental DNA metabarcoding, as a non-invasive, fast, universally applicable approach, is nowadays claiming the attention of researchers, stakeholders and governmental agencies. Therefore, exploring, evaluating and finally establishing the quantitative value of such a broadly-used tool for diversity monitoring is essential.

Data Resources

Details of protocols, bioinformatics, R script and supplementary material used for the analyses can be found on Open Science Framework at: DOI https://doi.org/10.17605/OSF.IO/ZWPSQ. Sequencing data have been submitted to NCBI (Bioproject: PRJNA638011; SRA accession numbers: SRR11949830–SRR11949928).

References

Alberdi A, Ostaizka A, Gilbert MTP, Bohmann K (2018) Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution 9(1): 134–147. https://doi.org/10.1111/2041-210X.12849

Barnes MA, Turner CR, Jerde CL, Renshaw MA, Chadderton WL, Lodge DM (2014) Environmental conditions influence eDNA persistence in aquatic systems. Environmental Science & Technology 48(3): 1819–1827. https://doi.org/10.1021/es404734p

Biggs J, Ewald N, Valentini A, Gaboriaud C, Dejean T, Griffiths RA, Williams P (2015) Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). Biological Conservation 183: 19–28. https://doi.org/10.1016/j.biocon.2014.11.029

Blackman RC, Constable D, Hahn C, Sheard AM, Durkota J, Hänfling B, Handley LL (2017) Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples – first record of *Gammarus fossarum* in the UK. Aquatic Invasions 12(2): 177–189. https://doi.org/10.3391/ai.2017.12.2.06

Britton JR, Brazier M, Davies GD, Chare SI (2008) Case studies on eradicating the Asiatic cyprinid *Pseudorasbora parva* from fishing lakes in England to prevent their riverine dispersal. Aquatic Conservation: Marine and Freshwater Ecosystems 18(6): 867–876. https://doi.org/10.1002/aqc.919

Bronner IF, Quail MA, Turner DJ, Swerdlow H (2013) Improved protocols for illumina sequencing. Current Protocols Human Genetics 79(1): 18.2.1–18.2.42 https://doi.org/10.1002/0471142905.hg1802s79

Buxton AS, Groombridge JJ, Griffiths RA (2018) Comparison of two citizen scientist methods for collecting pond water samples for environmental DNA studies. Citizen Science: Theory and Practice 3(2): e2. [9 pp.] https://doi.org/10.5334/cstp.151

Bylemans J, Gleeson DM, Duncan RP, Hardy CM, Furlan EM (2019) A performance evaluation of targeted eDNA and eDNA metabarcoding analyses for freshwater fishes. Environmental DNA 1(4): 402–414. https://doi.org/10.1002/edn3.41

Civade R, Dejean T, Valentini A, Roset N, Raymond J-C, Bonin A, Taberlet P, Pont D (2016) Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. PLoS ONE 11(6): e0157366. https://doi.org/10.1371/journal.pone.0157366

Collins RA, Bakker J, Wangensteen OS, Soto AZ, Corrigan L, Sims DW, Genner MJ, Mariani S. (2019) Non-specific amplification compromises environmental DNA metabarcoding with COI. Methods in Ecology and Evolution 10(11): 1985–2001. https://doi.org/10.1111/2041-210X.13276

Copp GH, Wesley KJ, Verreycken H, Russell IC (2007) When an 'invasive' fish species fails to invade! Example of the topmouth gudgeon *Pseudorasbora parva*. Aquatic Invasions 2: 107–112. https://doi.org/10.3391/ai.2007.2.2.4

Davison PI, Copp GH, Creach V, Vilizzi L, Britton JR (2017) Applications of environmental DNA analysis to inform invasive fish eradication operations. The Science of Nature 104: e35. https://doi.org/10.1007/s00114-017-1453-9

Davison PI, Falcou-Préfol M, Copp GH, Davies GD, Vilizzi L, Créach V (2019) Is it absent or is it present? Detection of a non-native fish to inform management decisions using a new highly-sensitive eDNA protocol. Biological Invasions 21(8): 2549–2560. https://doi.org/10.1007/s10530-019-01993-z

De Barba M, Miquel C, Boyer F, Mercier C, Rioux D, Coissac E, Taberlet P (2014) DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. Molecular Ecology Resources 14(2): 306–323. https://doi.org/10.1111/1755-0998.12188

Deiner K, Walser JC, Mächler E, Altermatt F (2015) Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. Biological Conservation 183: 53–63. https://doi.org/10.1016/j.biocon.2014.11.018

Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Pfrender ME (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology 26(21): 5872–5895. https://doi.org/10.1111/mec.14350

de Souza LS, Godwin JC, Renshaw MA, Larson E (2016) Environmental DNA (eDNA) detection probability is influenced by seasonal activity of organisms. PLoS ONE 11(10): e0165273. https://doi.org/10.1371/journal.pone.0165273

Djurhuus A, Port J, Closek CJ, Yamahara KM, Romero-Maraccini O, Walz KR, Chavez FP (2017) Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. Frontiers in Marine Science 4: e314. https://doi.org/10.3389/fmars.2017.00314

Eichmiller JJ, Miller LM, Sorensen PW (2016) Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. Molecular Ecology Resources 16(1): 56–68. https://doi.org/10.1111/1755-0998.12421

Evans NT, Olds BP, Renshaw MA, Turner CR, Li, Y, Jerde CL, Lodge DM (2016) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. Molecular Ecology Resources 16(1): 29–41. https://doi.org/10.1111/1755-0998.12433

Evans NT, Shirey PD, Wieringa JG, Mahon AR, Lamberti GA (2017) Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. Fisheries 42(2): 90–99. https://doi.org/10.1080/03632415.2017.1276329

Ficetola GF, Romano A, Salvidio S, Sindaco R (2018) Optimizing monitoring schemes to detect trends in abundance over broad scales. Animal Conservation 21(3): 221–231. https://doi.org/10.1111/acv.12356

Fonseca VG (2018) Pitfalls in relative abundance estimation using eDNA metabarcoding. Molecular Ecology Resources 18(5): 923–926. https://doi.org/10.1111/1755-0998.12902

GB Non-Native Species Secretariat (2020) GB Non-Native Species Secretariat. www.nonnativespecies.org [Accessed May 2020]

Genovesi P, Carnevali L (2011) Island invasives: Eradication and management. Gland, Switzerland: IUCN.

Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Laramie MB (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution 7(11): 1299–1307. https://doi.org/10.1111/2041-210X.12595

Grey EK, Bernatchez L, Cassey P, Deiner K, Deveney M, Howland KL, Pfrender ME (2018) Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. Scientific Reports 8(1): 1–10. https://doi.org/10.1038/s41598-018-27048-2

Hänfling B, Lawson Handley L, Read DS, Hahn C, Li J, Nichols P, Winfield IJ (2016) Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. Molecular Ecology 25(13): 3101–3119. https://doi.org/10.1111/mec.13660

Harper LR, Buxton AS, Rees HC, Bruce K, Brys R, Halfmaerten D, Priestley V (2019) Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. Hydrobiologia 826(1): 25–41. https://doi.org/10.1007/s10750-018-3750-5

Hering D, Borja A, Jones JI, Pont D, Boets P, Bouchez A, Leese F (2018) Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. Water Resources 1(138): 192–205. https://doi.org/10.1016/j.watres.2018.03.003

Jerde CL (2019) Can We Manage Fisheries with the Inherent Uncertainty from eDNA? Journal of Fish Biology 187: 1–13. https://doi.org/10.1111/jfb.14218

Kassambara A (2018) ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.2. https://CRAN.R-project.org/package=ggpubr

Kelly RP, Port JA, Yamahara KM, Crowder LB (2014) Using environmental DNA to census marine fishes in a large mesocosm. PloS ONE 9(1): e86175. https://doi.org/10.1371/journal.pone.0086175

Kelly RP, Shelton AO, Gallego R (2019) Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. Scientific Reports 9(1): 1–14. https://doi.org/10.1038/s41598-019-48546-x

Klymus KE, Richter CA, Chapman DC, Paukert C (2015) Quantification of eDNA shedding rates from invasive bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix*. Biological Conservation 183: 77–84. https://doi.org/10.1016/j.biocon.2014.11.020

Kull T, Sammul M, Kull K, Lanno K, Tali K, Gruber B, Henle K (2008) Necessity and reality of monitoring threatened European vascular plants. Biodiversity Conservation 17(14): e3383. https://doi.org/10.1007/s10531-008-9432-2

Lacoursière-Roussel A, Rosabal M, Bernatchez L (2016) Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. Molecular Ecology Resources 16(6): 1401–1414. https://doi.org/10.1111/1755-0998.12522

Larson ER, Graham BM, Achury R, Coon JJ, Daniels MK, Gambrell DK, Reed EM (2020) From eDNA to citizen science: emerging tools for the early detection of invasive species. Frontiers in Ecology and Environ. 18(4): 194–202. https://doi.org/10.1002/fee.2162

Majaneva M, Diserud OH, Eagle SH, Boström E, Hajibabaei M, Ekrem T (2018) Environmental DNA filtration techniques affect recovered biodiversity. Scientific Reports 8(1): e4682. https://doi.org/10.1038/s41598-018-23052-8

Lamb PD, Hunter E, Pinnegar JK, Creer S, Davies RG, Taylor MI (2019) How quantitative is metabarcoding: A meta-analytical approach. Molecular Ecology 28(2): 420–430. https://doi.org/10.1111/mec.14920

Lawson Handley L, Read DS, Winfield IJ, Kimbell H, Johnson H, Li J, Szitenberg A (2019) Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. Environmental DNA 1(1): 26–39. https://doi.org/10.1002/edn3.5

Li J, Hatton-Ellis TW, Lawson Handley LJ, Kimbell HS, Benucci M, Peirson G, Hänfling B (2019a). Ground-truthing of a fish-based environmental DNA metabarcoding method for assessing the quality of lakes. Journal of Applied Ecology 56(5): 1232–1244. https://doi.org/10.1111/1365-2664.13352

Li J, Lawson Handley LJ, Harper LR, Brys R, Watson HV, Di Muri C, Hänfling B (2019b). Limited dispersion and quick degradation of environmental DNA in fish ponds inferred by metabarcoding. Environmental DNA 1(3): 238–250. https://doi.org/10.1002/edn3.24

Li J, Lawson Handley LJ, Read DS, Hänfling B (2018) The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. Molecular Ecology Resources 18(5): 1102–1114. https://doi.org/10.1111/1755-0998.12899

McKee AM, Spear SF, Pierson TW (2015) The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. Biological Conservation 183: 70–76. https://doi.org/10.1016/j.biocon.2014.11.031

Minamoto T, Naka T, Moji K, Maruyama A (2016) Techniques for the practical collection of environmental DNA: filter selection, preservation, and extraction. Limnology 17(1): 23–32. https://doi.org/10.1007/s10201-015-0457-4

Oksanen J, Blanchet FG, Kindt R, Legendre P, O'hara RB, Simpson GL, Wagner H (2010) Vegan: community ecology package. R package version 1: 17–4.

OS Open Map (2019) OS Open Map – Local [SHAPE geospatial data], Scale 1:10000, Tiles: ny,sd, Updated: 11 October 2019, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service. https://digimap.edina.ac.uk [Downloaded: 2020-02-14 10:00:11.794]

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Radinger J, Britton JR, Carlson SM, Magurran AE, Alcaraz-Hernández JD, Almodóvar A, Benejam L, Fernández-Delgado C, Nicola GG, Oliva-Paterna FJ, Torralva M (2019) Effective monitoring of freshwater fish. Fish and Fisheries 20: 729–747. https://doi.org/10.1111/faf.12373

Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Research 39(21): e145–e145. https://doi.org/10.1093/nar/gkr732

Robinson CV, Garcia de Leaniz C, Rolla M, Consuegra S (2019) Monitoring the eradication of the highly invasive topmouth gudgeon (Pseudorasbora parva) using a novel eDNA assay. Environmental DNA 1(1): 74–85. https://doi.org/10.1002/edn3.12

Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation 17: e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Sard NM, Herbst SJ, Nathan L, Uhrig G, Kanefsky J, Robinson JD, Scribner KT (2019) Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. Environmental DNA 1(4): 368–384. https://doi.org/10.1002/edn3.38

Sassoubre LM, Yamahara KM, Gardner LD, Block BA, Boehm AB (2016) Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. Environmental Science & Technology 50(19): 10456–10464. https://doi.org/10.1021/acs.est.6b03114

Sellers GS, Di Muri C, Gómez A, Hänfling B (2018) Mu-DNA: a modular universal DNA extraction method adaptable for a wide range of sample types. Metabarcoding & Metagenomics 2: e24556. https://doi.org/10.3897/mbmg.2.24556

Seymour M, Durance I, Cosby BJ, Ransom-Jones E, Deiner K, Ormerod SJ, Edwards F (2018) Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. Communications Biology 1(1): 1–8. https://doi.org/10.1038/s42003-017-0005-3

Sigsgaard EE, Torquato F, Frøslev TG, Moore AB, Sørensen JM, Range P, Thomsen PF (2019) Using vertebrate environmental DNA from seawater in biomonitoring of marine habitats. Conservation Biology 34(3): 697–710. https://doi.org/10.1111/cobi.13437

Spens J, Evans AR, Halfmaerten D, Knudsen SW, Sengupta ME, Mak SS, Hellström M (2017) Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. Methods in Ecology and Evolution 8(5): 635–645. https://doi.org/10.1111/2041-210X.12683

Takahara T, Minamoto T, Yamanaka H, Doi H, Kawabata ZI (2012) Estimation of fish biomass using environmental DNA. PloS ONE 7(4): e35868. https://doi.org/10.1371/journal.pone.0035868

Takahashi S, Sakata MK, Minamoto T, Masuda R (2020) Comparing the efficiency of open and enclosed filtration systems in environmental DNA quantification for fish and jellyfish. PLoS ONE 15(4): e0231718. https://doi.org/10.1371/journal.pone.0231718

Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M, Willerslev E (2012) Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS ONE 7(8): e41732. https://doi.org/10.1371/journal.pone.0041732

Turner CR, Uy KL, Everhart RC (2015) Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation 183: 93–102. https://doi.org/10.1016/j.biocon.2014.11.017

Ushio M, Murakami H, Masuda R, Sado T, Miya M, Sakurai S, Yamanaka H, Minamoto T, Kondoh M (2018) Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. Metabarcoding and Metagenomics 2: e23297. https://doi.org/10.1101/113472

Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, Gaboriaud C (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology 25(4): 929–942. https://doi.org/10.1111/mec.13428

Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer, 260 pp. https://doi.org/10.1007/978-3-319-24277-4

Winfield IJ, Fletcher JM, James JB, Bean CW (2009) Assessment of fish populations in still waters using hydroacoustics and survey gill netting: experiences with Arctic charr (Salvelinus alpinus) in the UK. Fisheries Research 96(1): 30–38. https://doi.org/10.1016/j.fishres.2008.09.013

Wood SA, Pochon X, Laroche O, von Ammon U, Adamson J, Zaiko A (2019) A comparison of droplet digital polymerase chain reaction (PCR), quantitative PCR and metabarcoding for species-specific detection in environmental DNA. Molecular Ecology Resources 19(6): 1407–1419. https://doi.org/10.1111/1755-0998.13055

Zhang S, Lu Q, Wang Y, Wang X, Zhao J, Yao M (2019) Assessment of fish communities using environmental DNA: Effect of spatial sampling design in lentic systems of different sizes. Molecular Ecology Resources 20(1): 242–255. https://doi.org/10.1111/1755-0998.13105

Supplementary material 1
Table S1 and Figure S1
Authors: Cristina Di Muri, Lori Lawson Handley, Colin W. Bean, Jianlong Li, Graeme Peirson, Graham S. Sellers, Kerry Walsh, Hayley V. Watson, Ian J. Winfield, Bernd Hänfling
Data type: Protocols, figures and table
Explanation note: Additional laboratory protocol information and fish diversity and biomass/abundance data of the ponds surveyed; Figure S1 – Read counts and site occupancy barplots; Table S1 – Fish biomass/abundance data; Protocol of DNA extraction from Sterivex filters; Magnetic beads purification protocol.

Link: https://doi.org/10.3897/mbmg.4.56959.suppl1

Supplementary material 2
Table S2. Fish taxonomic assignment metaBEAT

Authors: Cristina Di Muri, Lori Lawson Handley, Colin W. Bean, Jianlong Li, Graeme Peirson, Graham S. Sellers, Kerry Walsh, Hayley V. Watson, Ian J. Winfield, Bernd Hänfling

Data type: Table

Explanation note: metaBEAT results of taxonomic assignment (100% identity) against the custom-curated 12S fish database.

Link: https://doi.org/10.3897/mbmg.4.56959.suppl2

Supplementary material 3
Table S3. Unassigned blast 1.0

Authors: Cristina Di Muri, Lori Lawson Handley, Colin W. Bean, Jianlong Li, Graeme Peirson, Graham S. Sellers, Kerry Walsh, Hayley V. Watson, Ian J. Winfield, Bernd Hänfling

Data type: Table

Explanation note: Results of unassigned reads blast (100% identity) against GenBank.

Link: https://doi.org/10.3897/mbmg.4.56959.suppl3